University of Massachusetts Amherst

ScholarWorks@UMass Amherst

International CHRIE Conference-Refereed Track

2010 ICHRIE Conference

Jul 31st, 9:45 AM - 10:45 AM

The Diversity-Validity Dilemma: Implications from the Ricci v. DeStefano U.S. Supreme Court Case

Juan M. Madera
University of Houston - Main, jmmadera@uh.edu

JeAnna Abbott
University of Houston - Main, jabbott@uh.edu

The Diversity-Validity Dilemma: Implications from the Ricci v. DeStefano U.S. Supreme Court Case

Juan M. Madera
University of Houston, Conrad N. Hilton College
Houston, Texas, USA

and

JeAnna Abbott
University of Houston, Conrad N. Hilton College
Houston, Texas, USA

ABSTRACT

One of the most critical problems that hospitality firms face in selecting employees is that the use of cognitive ability tests and other valid predictors of job performance lead to subgroup differences between majority and minority group members. The recent Ricci v. DeStefano U.S. Supreme Court case provides new implications for the adverse diversity-validity dilemma. This paper seeks to bridge areas that have not been discussed in unison, including the costs of employment discrimination claims, the validity-diversity dilemma, and disparate treatment v. disparate impact under the Supreme Court's recent ruling, and alternative selection methods.

Keywords: Diversity, Validity, Discrimination, Selection, Testing

INTRODUCTION

The demographic composition of the United States is rapidly changing. By 2020, the population of the United States is expected to increase by 42 million (U.S. Bureau of Labor Statistics, 2009). It is estimated that Hispanics will compose 47%, African Americans 22%, Asians and other ethnic minority groups 18%, and Whites 13% of this change (Chen, 1994). The U.S. Bureau of Labor Statistics (2009) estimates that by 2020, Hispanics will compose 17%, African Americans 14%, Asians and other ethnic minority groups 7%, and Whites 79.2% of the workforce (U.S. Bureau of Labor Statistics, 2009). How do these changes impact organizations? According to the attraction theory (Byrne, 1971), the organization that "looks like" the customers it serves will be more responsive to these customers' needs (Auh, 2005). Accordingly, hospitality companies must give more thought to how they attract minority customers to their services. Changing work-force demographics have also led organizations to reexamine practices for attracting new workers. Recruitment strategies, once developed for a homogeneous labor force, must now be reconsidered (Thomas & Wise, 1999). In essence, those organizations, especially hospitality firms that are able to embrace these "new" components of the labor market (race, ethnicity and sex based), will be able to take advantage of the full pool of potential employees. Not only will their employee base be stronger, but these organizations will be able to respond more effectively to changing target population needs. Thus, recruiting, selecting and retaining a diverse work force will give the organization a competitive advantage (Richard, 2000; Pitts, 2009).

Arguably, all organizations want to hire or promote the best candidates for a particular job. This usually means selecting individuals that are both well-suited and have the ability to perform the particular job. Thus, many organizations use selection tools (e.g. tests of cognitive ability) that predict high-quality candidates. Although many selection tools have been found to be good predictors of job performance, they also produce varying degrees of subgroup differences (Pyburn, Ployhart, & Kravitz, 2008). In particular, tests of cognitive ability are the best predictors of job performance, but also produce large subgroup differences between majority and minority group members (Outtz, 2002; Sackett, Schmitt, Ellingson, & Kabin, 2001). In other words, ethnic minority groups tend to

score lower than majority groups. These subgroup differences in test performance can cause disparate impact. There are many reasons why an organization might want to avoid disparate impact in selection. Yet, foremost on most employers' minds is the fear of a lawsuit. The desire to hire high-quality candidates, increase diversity and avoid litigation seems to be in conflict. This paper will address the tension between selection methods, test validity and claims of discrimination (i.e. disparate impact). In order to accomplish this, it is necessary to bridge a few areas that have not been discussed in unison, including the costs of employment discrimination claims, the validity-diversity dilemma, and disparate treatment v. disparate impact under the Supreme Court's recent ruling. We will also discuss alternative selection methods that are valid predictors of job performance and may be more appropriate to use in the hospitality industry context.

Costs of Employment Discrimination Claims

Racio-ethnic and sex based discrimination claims can lead to legal action, including administrative proceedings and litigation, resulting in significant legal costs. In this regard, the U.S. Equal Employment Opportunity Commission (EEOC) received 72,741 claims of racio-ethnic or sex based discrimination in 2009 alone, and it obtained 229.6 million in monetary benefits for victims (Equal Employment Opportunity Commission, 2009). Of these claims, 67,391 were resolved by the EEOC (Equal Employment Opportunity Commission, 2009). According to national data from Jury Verdict Research, the average cost to defend an employment practices lawsuit, including discrimination claims is in excess of $100,000 for a single plaintiff and class action lawsuits can be well over $3,000,000 (Lipton, 2009). And, if the judgment is in favor of the complainant, it is not uncommon for the award to be devastating. For example, recent settlements involving Coca-Cola ($192 million), Publix Markets ($81 million) and Shoney's ($105 million) were all quite large (Selmi, 2003). Additional costs may be incurred in increased insurance rates, decreases in stock prices, and damage to the organization's reputation if there is publicity surrounding the claim (Goldman, Gutek, Stein, & Lewis, 2006). A study conducted by Wright, Ferris, Hiller, and Kroll (1995) compared companies with exemplary diversity management practices with those that paid legal damages to settle discrimination lawsuits. The results show that the exemplary firms perform better, as measured by their stock prices. Again, it is clear that there are many benefits to recruiting, selecting and retaining a diverse workforce. What is less clear, however, is how hospitality firms can recruit, attract, and manage a diverse workforce effectively.

Diversity-validity dilemma

Obviously, an organization wants to use selection tools that provide the best predictors for high quality candidates. This means identifying individuals that have the knowledge, skills, abilities and other characteristics (KSAOs) most relevant to individual job performance (Pyburn et al., 2008). Unfortunately, the scores of non-whites and women on predictor scores in many selection devices are problematic. In other words, minorities do not perform as well on many of the selection tools that tend to be the best predictors of job performance. Thus, the organization winds up with a trade-off between the validity of the selection procedure and the goals of hiring a diverse workforce. This problem is referred to as the diversity-validity dilemma. Since organizations are not required by law to use more valid selection tools, many organizations have responded to this dilemma by simply using less valid selection tools. These selection tools generally do not have lower pass rates for minorities. Nonetheless, the less valid selection tools also do not predict job performance as well. Another benefit to the organization of using a less valid selection tool is avoiding the threat of being sued for disparate impact. However, recently, the United States Supreme Court in Ricci v. DeStanfano, 229 S. Ct. 1158 (2009) was forced to decide the question: Can the threat of a disparate impact claim create a basis for the organization to engage in discriminatory treatment?

Summary of the Facts in Ricci v. DeStefano

In Ricci v. DeStefano (2009), the city hired a company to design and administer a job-related, statistically valid test that would be given to firefighters who wanted to be promoted. Despite the attempts to ensure that the test

would not favor White candidates, the test results produced a statistically significant disparity between the more successful White candidates and the not-as-successful African-American and Hispanic candidates. Under EEOC rules, the disparity alone created the legal presumption of disparate impact. Both sides threatened to sue the city. Ultimately, the city discarded the test results relying on the race-based statistics. The White firefighters sued the city and lost in both lower court decisions.

Test Validity and the Ricci Decision

At the center of the Ricci case is discrimination. Thus, in order to understand the Ricci case it is necessary to explain the difference between disparate treatment and disparate impact claims. Disparate treatment is intentional discrimination. In a disparate treatment case, liability depends on whether the protected trait actually motivated the employer's decision. In respect to the Ricci case, disparate treatment occurred when the city did not certify the results of the test. In doing so, the city intentionally did not promote the seventeen incumbent White firefighters and one Hispanic firefighter that passed the promotion test. The reason the city did not certify the result of the test was that incumbent African-American firefighters' scores did not pass the score for promotion. Consequently, the city officials feared that the African-American firefighters might sue the city because of disparate impact.

Disparate impact occurs when a test, procedure, or standard is applied to all employees, but they lead to a substantial difference in employment outcomes (e.g., selection, promotion) for members of a particular group *and* if they are invalid—unrelated to job performance or there is an equally valid test that is less discriminatory. For example, a minimum height requirement of 5' 10" for firefighters would have adverse impact on women, who on average are shorter than men. To use it, an employer must show that the required height is related to successful job performance (i.e., evidence for validity of the requirement).

Thus, the validity of the promotion test used by the fire department was a key issue. Valid tests are the backbone of fair selection and promotion decisions—without validity, employers would make selection and promotion decisions that are not related to job performance. The city provided no concrete evidence that the test was invalid or unrelated to the firefighter position. This was a vital step for the city to demonstrate, because not certifying the results of the test that have substantial differences for members of a particular group can be justified by strong basis in evidence of disparate impact. That is, disparate impact only exists if the test that led to the racial differences in the Ricci v. DeStefano case was invalid or if there was an equally valid, less discriminatory test. The Supreme Court sided with the White firefighters because the city did not show a "strong basis in evidence" that it would lose an adverse impact case to reject a promotional examination on these grounds**.** Justice Kennedy, who wrote the opinion for a 5-4 majority, asserted that the City of New Haven could be liable of disparate impact only if the test was not job-related or there is an equally valid test that is less discriminatory that the city could have used. Essentially, the court held that a statistical disparity alone can establish a prima facie case of disparate impact.

However, statistical disparity alone cannot constitute a strong basis in the evidence for an employer's rejection of test scores. Thus, employers must do more than determine that hiring and promotion tests scores create a statistical disparity. As such, the Supreme Court answered the question: Can an employer use the threat of being sued for disparate impact as a legal defense for intentional discrimination? No.

In light of the diversity–validity dilemma, disparate treatment, and disparate impact discrimination—which were at the center of the Ricci v. DeStefano case—researchers have examined alternative methods so that organizations can use tests that effectively predict future performance, but also minimize differences across ethnic groups (Ployhart & Holtz, 2008; Pyburn, Ployhart, & Kravitz, 2008). There are two areas of research that provide alternative solutions for the diversity–validity dilemma. The first area of research involves the use of alternative predictor constructs, such as using measures of personality (e.g., measures of conscientious, extraversion, or agreeableness). The second area of research examines the use of alternative predictor measurement methods, such as

using interviews or assessment centers, which measure multiple constructs simultaneously (e.g., cognitive ability and personality).

Alternative Predictor Constructs

Alternative predictor constructs tap a single latent construct, such as conscientiousness or general cognitive ability, and are related to job performance. Using data from line-level restaurant employees, Tracey, Sturman, and Tews (2007) showed that general cognitive ability and conscientiousness were significant predictors of performance among restaurant employees. More importantly, however, conscientiousness was a better predictor of performance of experienced employees, whereas general cognitive ability was a better predictor of performance for new employees. As such, the authors showed that measures of conscientiousness can be significant and useful predictors of job performance for hospitality firms. Conscientiousness is not the only valid personality predictor of job performance; other personality factors include extraversion, agreeableness, open to experience, and emotional stability (Barrick & Mount, 1991).

Barrick and Mount's (1991) meta-analysis showed that conscientiousness is the best predictor of overall job performance across job types and industries; extraversion is a valid predictor for job performance in management and sales; extraversion and openness to experience were valid predictors of training proficiency criteria; and extraversion and agreeableness valid predictors for performance of front-line employees. As such, these personality constructs are useful alternative predictors that firms can use to help select employees.

Not only are these personality measures valid predictors of job performance, these measures also produce smaller differences across ethnic groups (Ployart & Hotz, 2008). For example, Ployart and Holtz (2008) reported group differences using the d statistic (the mean of the majority [i.e., Caucasian] group minus the mean of ethnic minority group [i.e., African-American] divided by their pooled standard deviations). General cognitive ability produced a d value of .99, which indicates that Caucasian and ethnic minorities differ by .99 standard deviation units, representing a large difference. In contrast, the d value for conscientiousness was .06, representing a small difference between Caucasian and ethnic minority employees.

One of the most stimulating ideas to emerge from organizational psychology is the concept of emotional intelligence, which provides another alternative predictor construct (Cote & Miners, 2006). Emotional intelligence is proposed as the ability to perceive one's emotions and the emotions of others, regulate emotions in the self and others, and use emotions to facilitate performance (Mayer, Caruso, & Salovey, 1999). Research shows that emotional intelligence significantly predicts job performance and that the relationship between emotional intelligence and job performance is stronger as cognitive ability decreases (Cote & Miners, 2006). That is, emotional intelligence can be compensatory; employees with low cognitive intelligence can perform effectively if they are emotionally intelligent.

More importantly, research suggests that emotional intelligence can minimize differences in cognitive ability scores across ethnic groups. In particular, African-Americans scored one-third of a standard deviation (d = 0.32) above Caucasians on emotional intelligence (Van Rooy, Alexander, & Chockalingam, 2005). Thus, emotional intelligence tests can favor minority groups, mitigating potential adverse impact concerns—though, more research is needed in this area.

While this research offers alternative predictors that are valid predictors of job performance and produce smaller differences across ethnic groups, these measures are not better predictors than cognitive ability. Thus, we take caution that using personality measures will result in better selection results than using measures of cognitive ability. Instead, using personality measures with measures of cognitive ability might be more beneficial for employers to use than using personality or cognitive ability measures alone.

Alternative Predictor Measurement Methods

In addition to examining the relationships between predictor constructs (e.g., cognitive ability and personality) and job performance, researchers have also examined alternative methods to measure such constructs. Alternative predictor measurement methods, such as using interviews or assessment centers, measure multiple constructs simultaneously.

Structured interviews are one of the most effective predictor measurement methods. Research shows that structured interviews can be as valid and reliable predictors of job performance as are cognitive ability tests (Campion, Campion, & Hudson, 1994; Campion, Palmer, & Campion, 1997). In fact, Huffcutt and Arthur (1994) reported predictive validities for structured interviews from .35 to .62 for job performance among entry-level employees. Predictive validity is a measure of how well a variable or a set of variables predicts an outcome, such as job performance. Schmidt and Hunter (1998) reported a high predictive validity of .52, which was as large as the validity for cognitive ability .51.

More importantly, structured interviews (i.e., using predetermined questions for every applicant) produce an assessment of job candidates that is less open to interviewer bias (Campion et al., 1997; Huffcutt & Roth, 1998). The difference in structured interview scores between Caucasian and African-American applicants have been reported to be significantly lower (i.e., d value of .23) than the differences in cognitive ability tests (d = .99; Bobko, Roth, & Potosky, 1999; Ployart & Holtz, 2008). As such, structured interviews are less likely to lead to adverse impact. In fact, Williamson, Campion, Malos, Roehling, and Campion (1997) linked interview structure and litigation outcomes conceptually and empirically, and found that structure enhances interview reliability and validity. Structure is also linked to litigation outcomes, because a) structure increases consistency across candidates, which helps justify business necessity, b) structure reduces the overall subjectivity and, therefore, potential for bias in the decision-making process, c) and many aspects of structure are likely to enhance perceptions of procedural justice. Despite the positive results of structured interviews, a major limitation of structured interviews is that ethnic subgroup differences increases as the cognitive loading of the interview increases.

Another significant predictor measurement method that provides a solution for the diversity-validity dilemma is the assessment center, which is an array of standardized tests (e.g., job-related simulations, interviews, and/or psychological tests). Like the structured interview, assessment centers display high predictive validity, but also produce little adverse impact (Ployart & Hotz, 2008). Ethnic group differences, however, increases as the cognitive load of the tests increases (Goldstein, Yusko, & Nicolopoulos, 2001).

Conclusion

The recent and highly publicized Ricci v. DeStefano U.S. Supreme Court case served as a medium for discussions and research on the diversity-dilemma phenomena. A potential implication from the Ricci v. DeStefano decision for the hospitality industry is that disparate treatment cannot be used to avoid disparate impact. That is, a hospitality firm cannot intentionally throw out test results when minorities do not pass the test cut-off scores to avoid a lawsuit. The decision in the Ricci case sets restrictions for when an employer can and cannot use the results of tests. This case also emphasizes the need for alternative predictor constructs and predictor methods. Personality tests, emotional intelligence tests, structured interviews, and assessment centers are just some alternatives that have research to support their validity. A common belief in the industry is that the best employees for hospitality jobs are those that have the right personality and attitude, and fortunately, data from empirical research substantiates this line of reasoning (Tracey et al., 2007).

REFERENCES

Auh, S. (2005). The effects of soft and hard service attributes on loyalty: The mediating role of trust. *The Journal of*

Service Marketing, 19, 81-92.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Bobko, P., Roth, P.L., & Potosky, D. (1999). Derivation and implications of a meta analysis matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology, 52*, 561–589.

U.S. Department of Labor (2009). *Civilian labor force by age, sex, race, and Hispanic origin* Bureau of Labor Statistics. Retrieved on April 22, 2010 from www.bls.gov.

Byrne, D. E., (1971). *The Attraction Paradigm,* Academic Press, New York, NY.

Campion, M. A., Campion, J. E., & Hudson, J. P. (1994). Structured interviewing: A note on incremental validity and alternative question types. *Journal of Applied Psychology, 79*, 998-1002.

Campion, M.A., Palmer, D.K., & Campion, J.E. (1997). A review of structure in the selection interview. *Personnel Psychology, 50,* 655-703.

Chen, C.C., & Eastman, W. (1994). Toward a corporate civic culture of racial diversity. Paper presented at the meeting of Context of Diversity Conference, Greensboro, NC.

Cote, S., & Miners, C. (2006). Emotional intelligence, cognitive intelligence, and job performance. *Administrative Science Quarterly, 51*, 1-28.

Goldstein, H.W., Yusko, K.P., Nicolopoulos, V. (2001). Exploring Black-White subgroup differences of managerial competencies. Personnel *Psychology, 54*, 783–807.

Huffcutt, A. I., & Arthur, W., Jr. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology, 79,* 184-190.

Huffcutt, A. I., & Roth, P. L. (1998). Racial group differences in employment interview evaluations. *Journal of Applied Psychology, 83,* 179-189.

Lipton, B. (2009). Employment practices liability insurance-EPLI insurance claims on the rise. *American Chronicle.*

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for emotional intelligence. *Intelligence, 27*, 267–298.

Outtz, J. L. (2002). The role of cognitive ability tests in employment selection. *Human Performance, 15,* 161–171.

Pitts, D. (2009). Diversity management, job satisfaction, and performance: Evidence from U.S. federal agencies. *Public Administration Review,* 328-338.

Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology, 61*, 153-172.

Pyburn, K. M., Jr., Ployhart, R.E., & Kravitz, D.A. (2008). The diversity-validity dilemma: Overview and legal context. *Personnel Psychology, 61*, 143–151.

Ricci v. DeStanfano, 229 S. Ct. 1158 (2009).

Richard, O. C. (2000). Racial diversity, business strategy, and firm performance: A resource-based view. *Academy of Management Journal, 43,* 164-177.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). Highstakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative action world. *American Psychologist, 56,* 302–318.

Schmidt, F.L., Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262–274.

Thomas, K.M., & Wise, P.G. (1999). Organizational attractiveness and individual differences: Are diverse applicants attracted by different factors. *Journal of Business and Psychology,* 13, 375-390.

Tracey, J. B., Sturman, M. C., & Tews, M. J. (2007). Ability versus personality: Factors that predict employee job performance. *Cornell Quarterly, 48*, 313-322.

Van Rooy, D. L. V., Alexander, A., & Chockalingam, V. (2005). Group differences in emotional intelligence scores: Theoretical and practical implication. *Personality and Individual Differences, 38*, 689-700.

Williamson, L. G., Campion, J. E., Malos, S. B., Roehling, M. V., & Campion, M. A. (1997). Employment interview on trial linking interview structure with litigation outcomes. *Journal of Applied Psychology, 82,* 900-912.